A MIR4646 associated methylation locus is hypomethylated in adolescent depression

Boström, Adrian E ; Ciuculete, Diana-Maria ; Attwood, Misty ; Krattinger, Regina ; Nikontovic, Lamia ; Titova, Olga E ; Kullak-Ublick, Gerd A ; Mwinyi, Jessica ; Schiöth, Helgi B

Abstract: BACKGROUND: Studies of epigenetics and transcriptional activity in adolescents may provide knowledge about possible preventive strategies of depression. METHODS: We present a methylome-wide association study (MWAS) and cohort validation analysis of depression in adolescents, in two separate cohorts: discovery (n=93) and validation data set 1 (n=78). The genome-wide methylation pattern was measured from whole blood using the Illumina 450K array. A second validation cohort, validation data set 2, consists of post-mortem brain biopsies from depressed adults (n=58). We performed a MWAS by robust multiple linear regressions of methylation to a modified risk-score assessment of depression. Methylation levels of candidate CpG sites were correlated with expression levels of the associated gene in an independent cohort of 11 healthy volunteers. RESULTS: The methylation state of two CpG sites reliably predicted ratings of depression in adolescents (cg13227623 and cg04102384) ($p<10E-06$). Cohort validation analysis confirmed cg04102384 - located in the promoter region of microRNA 4646 (MIR4646) - to be hypomethylated in both validation data set 1 and validation data set 2 ($p<0.05$). Cg04102384 was inversely correlated to expression levels of MIR4646-3p in healthy controls ($p<0.05$). LIMITATIONS: MIR4646 was not differentially expressed in a subset of samples with adolescent depression measured by qRT-PCR measurements. CONCLUSION: We identify a specific MIR4646 associated epigenetic risk site to be associated with depression in adolescents. Cg04102384 putatively regulates gene expression of MIR4646-3p. Target gene prediction and gene set overrepresentation analysis revealed involvement of this miRNA in fatty acid elongation, a process related to omega-3 fatty acids, previously associated with depression.

DOI: https://doi.org/10.1016/j.jad.2017.05.017

methylation locus is hypomethylated in adolescent depression. Journal of Affective Disorders, 220:117-128.
DOI: https://doi.org/10.1016/j.jad.2017.05.017

A MIR4646 associated methylation locus is hypomethylated in adolescent depression

Adrian E. Boström[a], Diana-Maria Ciuculete[a], Misty Attwood[a], Regina Krattinger[b], Lamia Nikontovic[a], Olga E. Titova[a], Gerd A. Kullack-Ublick[b], Jessica Mwinyi[a], Helgi B. Schiöth[a]

[a] Division of Pharmacology, Department of Neuroscience, Uppsala University, Uppsala, Sweden

[b] Department of Clinical Pharmacology and Toxicology, University Hospital Zurich, University of Zurich, Switzerland

University of Uppsala

Department of Neuroscience (Medical faculty)

Box 593, Husargatan 3, Uppsala, Sweden.

Email: helgi.schioth@neuro.uu.se

Phone: +46 18 471 4136

Fax: +46 18 51 1540

Abstract

Background: Studies of epigenetics and transcriptional activity in adolescents may provide knowledge about possible preventive strategies of depression.

Methods: We present a methylome-wide association study (MWAS) and cohort validation analysis of depression in adolescents, in two separate cohorts: discovery (n=93) and validation (n=78). The genome-wide methylation pattern was measured from whole blood using the Illumina 450K array. We performed a MWAS by robust multiple linear regressions of methylation to a modified risk-score assessment of depression, adjusting for optimally determined co-variates. Candidate CpG loci were further investigated in a separate cohort of 78 adolescents. Methylation levels of candidate CpG sites were correlated with expression levels of the associated gene in an independent cohort of 11 healthy volunteers.

Results: The methylation state of two CpG sites reliably predicted ratings of depression in adolescents (cg13227623 and cg04102384) ($p<10\text{-}06$). In a cohort validation analysis, cg04102384 – located in the promoter region of microRNA 4646 (MIR4646) - was confirmed to be hypomethylated in depressed subjects ($p<0.05$). We provide evidence for an inverse correlation between cg04102384 methylation levels and expression levels of MIR4646-3p ($p<0.05$) in healthy controls.

Limitations: MIR4646 was not differentially expressed in a subset of samples with adolescent depression measured by qRT-PCR measurements.

Conclusion: We identify and confirm a specific MIR4646 associated epigenetic risk site to be associated with depression in adolescents. Cg04102384 putatively regulates gene expression of MIR4646-3p. Target gene prediction and gene set overrepresentation analysis revealed involvement of this miRNA in fatty acid elongation, a process related to omega-3 fatty acids, previously associated with depression.

Keywords: MIR4646; methylome-wide association study; adolescent depression; omega-3; microRNA

Introduction

Depressive disorders are highly prevalent during the adolescence and are associated with a higher risk for suicide, antisocial behavior, substance abuse, significant impairment and an increase of functional disability (Glied & Pine, 2002). Major depression (MDD) is common in young people, affecting up to 5-6% of the adolescents (Costello, Erkanli, & Angold, 2006). Moreover, a reliable detection of early onset depression is of high importance as a solid diagnosis during adolescence would help to efficiently unveil people at risk for depression at later stages of life (Birmaher, Arbelaez, & Brent, 2002). However, the exact mechanism that underlies the high risk for depression during the youth is poorly understood. Investigating the relationship between the MDD risk as evaluated by the Development and Well-being Assessment (DAWBA), epigenetics mechanism and transcriptional activity in adolescents may provide knowledge of preventive strategies of depression among adults.

In recent years, there has been increasing interest to elucidate the role of epigenetic modifications in the pathogenesis of psychiatric disorders (D'Addario et al., 2012; Dell'Osso et al., 2014; Ruzicka, Subburaju, & Benes, 2015; Walker et al., 2016). In this context, especially changes in the methylation at CpG dinucleotides within regulatory regions of the DNA were studied, which were shown to be responsive to environmental signals by modifying the

transcription of genes (Cordova-Palomera et al., 2015). Notably, the importance of altered gene expression in MDD was recently demonstrated in a study where genes, such as *DVL3*, *CALM1* and *NMUR1* were differentially expressed in depressed individuals (Jansen et al., 2015), resulting in a new complementary mechanistic insight. Moreover, DNA methylation has been associated with other genomic functions, such as alternative splicing and promoter usage (Maunakea et al., 2010), which lead as well to a transcriptional modulation of MDD related genes.

Dysregulation of epigenetic control is particularly interesting in the context of MDD development (Menke & Binder, 2014), as the pathogenesis of this disease is characterized by a tight gene-environment interplay that influences neurobiological processes underlying the disease (Nugent, Tyrka, Carpenter, & Price, 2011; Schmitt, Malchow, Hasan, & Falkai, 2014). Importantly, depressive symptoms are thought to have an incidence spike during adolescence (Birmaher et al., 2002), suggesting a long-term importance for investigating the epigenetic changes during this time frame. Previous targeted studies suggested that the methylation profile of the *BDNF* gene could serve as biomarker of late-life depression (V Januar, Ancelin, Ritchie, Saffery, & Ryan, 2015). Furthermore, another study highlighted altered DNA methylation of the serotonin transporter (*5-HTT*) gene in depressed adolescents (Olsson et al., 2010). A recent genome-wide study identified differentially methylated loci between medication-free depressed individuals and non-psychiatric controls in blood, but the study was not corrected for cell-type heterogeneity (Numata et al., 2015), meriting further research in medication-free subject taking cell heterogeneity into account. The abundance of cell populations is relevant through their distinct signatures of DNA methylation (Reinius et al., 2012) that could induce false differences between depressed individuals and controls. Also, an epigenome-wide study reported differences of *PRIMA1* gene methylation between depressed adults and controls. However, the

authors were not able to replicate the findings in their additional samples (Sabunciyan et al., 2012).

Another attractive candidate mechanism to induce phenotype-associated changes is related to small non-coding RNAs (miRNA) which play a critical role in post-transcriptional fine-regulation of gene expression in many different tissues, including human brain functioning (Wei Chen & Qin, 2015). There is evidence suggesting an interaction between miRNA expression and psychiatric disorders, possibly mediated by DNA methylation (Villela et al., 2016). Due to the miRNAs role in degrading and/or inhibiting translation, a better understanding of the interaction between shifts in DNA methylation and a changed miRNA expression would provide additional valuable information about mechanistic links between methylation and depression.

Herein, we set out to identify the role of DNA methylation between whole-blood samples of adolescents at high risk for depression and controls in an unbiased methylome-wide approach, using the 450K Illumina array. The identified differentially methylated loci were further replicated in an independent cohort of depressed adolescents. We detect the gene encoding MIR4646 as significantly differentially methylated and investigate its change in expression further in an independent cohort of healthy controls. Lastly, we measure the miRNA levels for a sub-sample of the discovery data set and observe the expression change between adolescents at risk of depression and controls.

Methods

Discovery data set

The study of the Discovery data set was approved by the local ethics committee in Uppsala, Sweden (Regionala etikprövningsnämnden i Uppsala) and included 93 adolescents aged 14 to 17, recruited in the years 2013-2014. Both subjects and their participating parent(s) gave their written informed consent to participating in the study. Subjects were randomly selected from public school in Uppsala County and were included in the study if they exhibited an overall risk for psychiatric diagnoses of 15% or more, as measured by the DAWBA web-based diagnostic interview (described below (Goodman, Heiervang, Collishaw, & Goodman, 2011)). Self-reported information pertaining to basic physiological parameters and medication was provided by the participants. Body weight was measured for body mass index (BMI) calculation. The BMI z-scores were calculated and based on these values each subject was stratified into one of four weight category groups (underweight, normal weight, overweight and obese), as defined by the Center for Disease Control and Prevention (CDC). Subjects were grouped into three categories based on their DAWBA risk-score assessment for depression as ‘Controls' (~ 0-3%), '15% Risk Depression' (~ 15%) and 'Depressed' (~50-75%), respectively.

Replication data set

In the replication stage, 130 samples from the study mentioned above, but characterized and measured earlier in the time frame between November 2012 and January 2013, were studied. The study was approved by the local ethics committee in Uppsala, Sweden (Regionala etikprövningsnämnden i Uppsala) and all participants and their parent(s) gave their written informed consent. This data set is population-based and the same parameters were recorded as for the Discovery data set. In order to increase the power to detect meaningful differences, we

excluded subjects with intermediate risk scores for any psychiatric disorder, and stratified the remaining individuals into one of two categories: 'Controls' (~0%) and 'Depressed' (~15-50%).

Psychiatric Diagnoses

The DAWBA consists of web-based diagnostic interviews used to evaluate DSM-IV and ICD-10 type diagnoses specifically for individuals in the age range 5 to 17 years. The questionnaire is completed separately by both a parent and the child, and encompasses the most prevalent behavioral, emotional and hyperactivity type disorders. The covered categories include depression, generalized anxiety disorder, SAD, post-traumatic stress disorder (PTSD), autism and obsessive compulsive disorder (OCD). We used an algorithm to convert questionnaires to probability of diagnosis, ranging from less than 0.1% to 70% and reflecting the probability that an experienced clinical rater would assign the individual a corresponding DSM-IV or ICD-10 diagnosis (Goodman et al., 2011).

DNA Specimens

Body weight was measured and participants answered a questionnaire with questions about living conditions, place of birth and basic physiological parameters (height, age etc.). Venous blood was taken according to standard procedures and stored in six tubes at a total volume of 25 ml comprising two EDTA coated tubes for DNA extraction, two PAX gene tubes for RNA extraction, one lithium-heparin treated tube for plasma and one substrate-free tube for serum. All tubes were kept at minus 80°C after separation of plasma and serum by centrifugation.

Methylation Profiling

DNA was extracted using the phenolol-chloroform method (Sambrook, Fritsch, & Maniatis, 1989), and bisulfite converted by the EZ DNA Methylation - GoldTM kit (ZymoResearch, USA). Bisulfite converted DNA was hybridized to the Illumina 450k methylation chip

(Illumina, San Diego, CA, USA). The Illumina chip measures methylation at 485 777 CpG sites. The Illumina iScan system (Illumina, San Diego, CA, USA) was used for imaging of the array, whereby the methylation level of each CpG site was determined.

Data Preprocessing and Quality Control

Preprocessing of the methylation data was performed by adjustment of probe type differences, removal of batch effects and probe exclusion. Subsequently, principal component analysis (PCA) was used to identify sample outliers in the methylation data. Methylation preprocessing steps were performed in using R statistics (ww.r-project.org) together with the packages minfi (Aryee et al., 2014), wateRmelon (Schalkwyk, Pidsley, & Wong, 2013), ChAMP (Morris et al., 2014), sva (Leek, Johnson, Parker, Jaffe, & Storey, 2012), and limma (Smyth, 2004) of the Bioconductor project and the FactoMineR (Lê, Josse, & Mazet, 2008) package of the CRAN project. Concerning adjustment of probe type differences, removal of batch effects, probe exclusion and sample exclusion criteria please see Supplementary material.

Accounting for cellular heterogeneity

DNA methylation measured in whole blood is composed of different cell populations (Reinius et al., 2012). Rask-Andersen et al. demonstrated recently that changes in leukocyte fractions could introduce considerable variability in the DNA methylation pattern which could bias downstream analyses. Thus, it is important to take into account white blood cell type heterogeneity in genome-wide DNA methylation studies (Rask-Andersen et al., 2016). In the Discovery and Replication group, we implemented a minfi-based statistical procedure of the Houseman algorithm (Aryee et al., 2014), which uses raw intensity DNA methylation files to calculate the relative proportions of B cells, CD4+ and CD8+ T cells, granulocytes, monocytes, and natural killer cells.

Micro-RNA profiling

Sample inclusion considerations

Candidate microRNAs were investigated by measuring their expression levels in whole blood. For this purpose, we extracted RNA from a group of 27 subjects comprising eleven 'Controls', nine '15% risk depression' and seven 'Depressed' individuals from the Discovery data set. Samples were selected based on their level of DNA methylation at the associated methylation locus (cg04102384). In order to increase the power to detect meaningful associations, we selected the eleven controls with the lowest methylation levels, and 16 cases ('15% Risk Depression' or 'Depressed') with the highest methylation levels.

Reverse transcription and quantitative real-time PCR.

To measure miRNA expression, 10 ng of extracted RNA was reverse transcribed into cDNA using a TaqMan® miRNA Reverse Transcription Kit (Applied Biosystems, Waltham, MA, USA) and specific stem-loop reverse transcription primers (TaqMan® MicroRNA Assays MIR4646-3p and MIR4646-5p; Life Technologies, Carlsbad, CA, USA). RT-PCR was performed using 0.67 µl cDNA and 9.3 µl RT-PCR Universal Fast Master Mix (Applied Biosystems) including primers. U6snRNA/RNU44 was used as an internal control. All measurements were performed in triplicate.

Characterization of the expression data set

Eleven healthy male volunteers aged between 18 and 40 years were recruited from the region of Uppsala, Sweden, between 2013 and 2014. Blood analyses were performed before and after a meal intake. For the purpose of this study, only the non-fasting blood samples were further studied. The data was adjusted for white blood cell type heterogeneity. More details on the cohort and preprocessing of the methylation and RNA specimens have been previously published (Rask-Andersen et al., 2016).

Statistical analysis

All statistical analyses were performed in using R statistics, version 3.3.0.

Data analysis

Chi-squared tests with Monte Carlo computed p-values were used to detect differences in categorical variables, e.g. gender, BMI group (underweight, normal weight, overweight and obese), use of medications (sleep medications, neuroleptics, contraceptive pills, anxiolytics, ADHD medications and antidepressives) and the DAWBA risk score for any psychiatric disorder. In the Discovery group, ANOVA tests were used to investigate group differences in age. In the Replication group, which was stratified into two DAWBA groups, we used t-tests to investigate group differences in age.

Adjusting for potential confounders

In methylome-wide association studies, hidden confounders such as life style patterns or even prandial states can introduce unknown sources of bias. Drawing on a method for epigenome-wide analyses introduced by Zhagool et al. (Zaghlool et al., 2015), PCA-analysis was used to account for potential unmeasured sources of variation in the DNA methylation data using the MethylPCA tool (Wenan Chen et al., 2013). The known covariates (including age, gender, weight category, the relative proportions of white blood cell types and the DAWBA overall risk score for any psychiatric disorder) were regressed out prior to PCA. The calculated first ten principal components were considered as additional potential covariates in the methylome-wide analysis.

There were many potential covariates on the association analysis between DNA methylation and depression risk score, e.g. gender, age, self-reported use of medication, the white blood cell coefficients and the first ten principal components. To avoid overfitting by including too many covariates, we investigated each individual covariate against the phenotype of interest in

regression models using the 'lm' function in R. Covariates were incrementally and independently selected. Using the computed analysis of variance, we tested whether the addition of a particular covariate resulted in a better fit to the model and only included variables with a p-value < 0.05. The best linear model for depression risk score included the CpG sites, risk score for any psychiatric disorder ($p<0.00001$), self-reported use of ADHD medication ($p=0.022$), the relative proportion of CD4+ T cells ($p=0.017$) and the first principal component ($p=0.00081$). For statistical analysis, we transformed the beta values to M-values, which have been shown to be statistically more robust (Du et al., 2010).

Methylome-wide association study

The association between DNA methylation and depression was tested by linear models using the 'limma' package for R, applying an empirical Bayes method based on a moderated t-statistic (Smyth, 2004). We assumed a linear model where the M values of each CpG site were used as a quantitative dependent trait and the phenotype characterizing the risk for depression were used as covariates together with the other optimal covariates.

We used the package 'GenABEL' available for R to calculate the genomic inflation factor lambda for the epigenome-wide analysis (Aulchenko, Ripke, Isaacs, & van Duijn, 2007) to evaluate whether a general systemic inflation of significance values was abundant. In accordance to Zhagool et al. (Zaghlool et al., 2015), we first tested whether the calculated optimal covariates resulted in less systemic inflation as compared to limiting the covariates to age, gender and BMI z-score weight category. Using the optimal covariates in the methylome-wide association analysis between DNA methylation and depression risk score reduced the genomic inflation factor lambda from 1.95 to 1.11. To exclude any potential bias from systemic inflation of significance values (lambda > 1 in this case), all subsequent chi-squared statistics on a set of candidate markers were divided by lambda (Hinrichs, Larkin, & Suarez, 2009). For each CpG site, we thus divided the regression t-value by lambda, and used the adjusted t-values

to test for significance based on the t distribution. The inflation factor lambda for the adjusted p-values was estimated to be ~0.90, and systemic inflation of significance values was excluded. We used the bonferroni method correct for multiple testing. Bonferroni and lambda-adjusted p-values < 0.05 were considered significant.

Investigation of candidate CpG-sites in the Replication data set

In the Replication data set, we performed independent samples t-tests of candidate CpG-sites, contrasting methylation between cases and controls and taking the direction of the methylation change into account.

Correlation analyses between methylation and expression data

CpG sites consistently hyper- or hypomethylated in the Discovery and Replication data set were further investigated with regard to an association with transcriptional expression of the microRNA in focus using the Expression data set. We performed Spearman correlations of methylation M-values to normalized expression levels of probes associated with the candidate miRNA, assuming an inverse correlation. P-values < 0.05 were considered significant.

We also measured the miRNA levels of MIR4646-3p and MIR4646-5p by TaqMan analysis in 26 subjects from the Discovery data set and studied methylation-expression correlations by Spearman's rank correlation.

Differential expression analysis of MIR4646-3p/5p in a subset of samples from the Discovery data set

In the 26 subjects for whom miRNA levels were measured, we studied potential differences in normalized expression levels between the defined depression groups by ANOVA models, not taking any co-variates into account.

Functional analysis of the identified CpG site via chromatin states and long-range interactions

In order to illustrate the functional role of the identified CpG site in brain, as well as its potential regulatory effect on other genes, we performed chromatin states and long-range interactions analyses using the ENCODE project. The analysis uses Hidden Markov Models (HMMs), which were applied to seven brain regions, i.e. brain angular gyrus (BrainAG), brain anterior caudate (BrainAC), brain cingulate gyrus (BrainCG), brain hippocampus (BrainHIPPO), brain inferior temporal lobe (BrainITL), brain substantia nigra (BrainSN) and brain dorsolateral prefrontal cortex (BrainDPC), together with peripheral blood mononuclear primary cells (PBMC). As a result, an 18-state model was obtained, which, for simplicity, was reduced to five regions defining the relevant gene regulatory roles, which were indicated as (1) red, for active/flanking active/bivalent/poised transcription start site (TSS), (2) yellow, for active/bivalent/genic enhancer; orange, for flanking bivalent TSS/enhancer, green, for active transcription and grey, for repressed polyComb state (**Figure 3.**). The long-range interactions were investigated using chromatin analysis by paired-end tag sequencing (ChIA-PET). Several cell lines were used for the analysis, including erythrocytic leukaemia cells (K562), breast cancer (MCF-7), cervical cancer (HelaS3) and human colon carcinoma (HCT-116) cells, targeting the transcription factors RNA polymerase II and CTCF. Data was downloaded from the WashU Epigenome Browser, 37/hg19 version.

Target gene prediction and pathway analysis of MIR4646-3p

Candidate CpG-sites associated with microRNAs were further investigated by computationally predicting putative gene targets of the aforementioned miRNAs using the online webtool MiRWalk 2.0 (Dweep, Sticht, Pandey, & Gretz, 2011), a sophisticated online software tool that documents predictions from several independent prediction algorithms, including Targetscan

(Grimson et al., 2007), DIANA-microT-CDS (Paraskevopoulou et al., 2013), miRandarel2010 (Betel, Koppal, Agius, Sander, & Leslie, 2010) and RNAhybrid (Rehmsmeier, Steffen, Hochsmann, & Giegerich, 2004). MiRNA targets were considered as relevant hits when having a seed length ≥ 7 bases and when located within the 3′-UTR region. Genes identified as putative miRNA targets were further investigated by overrepresentation analysis of KEGG-defined pathways, using the online web tool 'ConsensusPathDB-human' (Kamburov, Stelzl, Lehrach, & Herwig, 2013).

Alignment and syntetic conservation across species

A thorough search of miRNA-4646 was performed in the miRBase database (release 21) using the webserver (Kozomara & Griffiths-Jones, 2014). A BLASTn (Boratyn et al., 2013) search was performed through the NCBI Blast webserver (v. 2.5.0) using the NCBI Genomic Reference Sequences dataset with word size 16 and all other parameters default values. The relevant hits were retained considering e-value (> 5e-06), sequence identity and appropriate sequence coverage, i.e., if the seed regions of the mature miRNAs were covered. The sequence region of interest for Homo sapiens (NC_000006.12) , Pan troglodytes (NC_006473.4), Pan paniscus (NW_014013975.1), Gorilla gorilla gorilla (NC_018430.1), Pongo abelii (NC_012597.1), Nomascus leucogenys (NC_019816.1), Macaca mulatta (NC_027896.1), Papio anubis (NW_003873063.1), Saimiri boliviensis (NW_003943814.1), Callithrix jacchus (NC_013899.1), Aotus nancymaae (NW_012186114.1), Mandrillus leucophaeus (NW_012106809.1), Rhinopithecus bieti (NW_016820117.1), Colobus angolensis palliates (NW_012115555.1), Cebus capucinus imitator (NW_016107339.1), Cercocebus atys (NW_012003394.1) were downloaded and then aligned using the Mafft webserver (v 7) (Katoh & Standley, 2013) with the L-INS-i method, allowing to adjust sequence direction, and all other settings default. Visualization with nucleotide coloring scheme and manual editing of the alignment was performed in Jalview (Waterhouse, Procter,

Martin, Clamp, & Barton, 2009) with further refinements performed in Adobe Illustrator. The Sequence-Structure Motif Base: pre-miRNA prediction webserver (http://www.regulatoryrna.org/webserver/SSMB/pre-miRNA/home.html) was used to assess if the identified genomic region in each investigated species was predicted to form pre-miRNA. The species phylogeny tree was created using phyloT (Letunic & Bork, 2016) and is based on NCBI taxonomy. The tree represents the evolutionary relationship among the investigated organisms; the branch lengths are not relative to evolutionary distances. The gene synteny was gleaned through the NCBI map view web portal (O'Leary et al., 2016).

Results

Behavior of the clinical outcome variables

In the Discovery data set, comprising 93 subjects and in the majority female, we initially aimed to identify CpG-sites, in which modifications of the epigenetic profile are associated with a modified risk score of depression. There were no significant differences between the three DAWBA sub-groups ('Controls', '15% Risk Depression' and 'Depressed') in age, gender, BMI z-score derived weight categories, in the use of medications or in the relative proportion of white blood cell type coefficients (CD4+ and CD8+ T cells, B cells, monocytes, NK cells, and granulocytes). As expected, the 'Depressed' subgroup had significantly higher DAWBA general risk score estimates for any psychiatric diagnosis ($p<0.001$). The Replication data set of 78 subjects included only female subjects. Depressed cases showed a higher BMI ($p<0.01$), took more often contraceptive pills ($p<0.01$) and scored higher in the DAWBA general risk score for any psychiatric diagnosis. There were no between-group differences in age, the relative proportion of white blood cell types or in the use of non-contraceptive medications (**Table 1.**).

Two CpG sites in proximity to ZWIM5 and MIR4646 are differentially methylated by depression risk score group

We performed multiple linear regression models of methylation M-values and depression risk group, to study the association between DNA methylation and a three-factorial risk score of depression in the Discovery data set, adjusting for the identified optimal covariates, comprising the DAWBA general risk score assessment of any psychiatric disorder, ADHD medication, relative proportion of CD4+ T-cells and the first principal component. This procedure effecticiently reduced the systemic inflation of significance values from 1.95 to 1.11, as evaluated by the lambda genomic inflation factor (**Figure 1a-1b.**). A potential bias through systemic inflation of significance values was successfully removed by adjusting the t-values for the lambda of 1.1, resulting in a new lambda of 0.9 (**Figure 1c.**). Additionally, Bonferroni-

correction was applied on the lambda-adjusted significance values. Two CpG sites were proven to be differentially methylated by depression risk score group (cg13227621 and cg04102384), associated with the genes *ZWIM5* and miRNA MIR4646 (p_{Bonf}<0.05)(**Table 2.**) (**Figure 2.**).

Methylation of cg04102384, which is associated with a modified DAWBA score of depression, is also hypomethylated in the replication data set

Hypomethylation of cg04102384, which was found to be associated with depression risk score in the Discovery data set, was confirmed to be hypomethylated in whole blood of the group of thirteen subjects with higher risk scores for depression as compared to the controls (replication data set)(**Table 3.**, p<0.05). Cg13227621 was not differentially methylated in relation to DAWBA.

CpG site methylation of cg04102384 is negatively correlated with expression of MIR4646-3p

To evaluate to what extent the methylation of cg04102384 is associated with the expression of the adjacent microRNA, Spearman correlations were performed using the Expression data set of eleven non-fasting healthy controls. The methylation state and the level of transcriptional miRNA expression were compared intra-individually to each other, assuming a negative correlation. Cg04102384 significantly inversely correlated with transcriptional levels of MIR4646-3p (p<0.05), but not with the transcriptional expression of MIR4646-5p or pre-miRNA levels of MIR4646.

To confirm these findings, we performed Taqman analysis, individually measuring the expression levels of MIR4646-3p and MIR4646-5p in 27 subjects from the Discovery data set (eleven 'Controls', nine '15% Risk Depression' and seven 'Depressed'). There were no between-group differences in expression levels of either of the two microRNAss. Nor could we

confirm an association between methylation and expression by Spearman's rank correlation method for either of the two miRNAs and the candidate CpG site.

A significant overrepresentation of has-miR-4646-3p putative gene targets are involved in biological processes associated with fatty acid elongation and biosynthesis of unsaturated fatty acids

Using the MirWalk2.0 analysis software(Dweep et al., 2011), we retrieved the predicted gene targets for has-miR-4646-3p. 552 genes were identified as putative gene targets for has-miR-4646-3p and were subsequently investigated by overrepresentation analysis of KEGG-defined pathways. There was a statistically significant overrepresentation of genes associated with fatty acid elongation (6 genes, q-value<0.01) and biosynthesis of unsaturated fatty acids (5 genes, q-value<0.05).

The CpG site next to the gene MIR4646 may have a regulatory role in brain and even modulate other genes

The investigation of the chromatin regions overlapping our CpG locus, cg04102384, revealed a regulatory role throughout all investigated brain regions, except brain cingulate gyrus. This CpG site is located in an enhancer region in blood, angular gyrus, dorsolateral prefrontal cortex, inferior temporal lobe and anterior caudate, and even in a TSS region in hippocampus and substantia nigra. Furthermore, the long-range interactions for this CpG site, represented by arcs (**Figure 3.**), refer to possible regulatory effects on multiple other genes, e.g. *CLIC1* and *BAG6*.

MIR4646 appears to be evolutionary conserved from the advent of Simiiformes (monkey)

The miRNA-4646 stem-loop region is identified only in Homo sapiens in the mirBase database and a BLASTn search in the NCBI webserver was used to investigate if this region is conserved in other species. Sixteen relevant hits were obtained with particular emphasis on the conservation of the seed regions. The conservation throughout the seed regions and the mature 5p and 3p regions are presented in the multiple sequence alignment (**Figure 4; Part A**). All sixteen sequences are predicted to fold into a pre-microRNA folding structure, i.e., hairpin loop. The gene synteny (**Figure 4; Part B**) is conserved among 13 of the 16 species investigated, except for variation in the number of LY6G6C/6D/6E/6F genes. Several of the species did not have an available mapped genome through the NCBI map view, or the mapping was incomplete. The reverse complement of the miRNA-4646-3p mature sequence had nine relevant hits with the miR-204 family when searched in the miRBase database. The conservation of the mature miRNA with the seed region emphasized is presented in the multiple sequence alignment (**Figure 5**).

Discussion

Using a methylome -wide study approach, we identified two genes, *ZSWIM5* and MIR4646, to be differentially methylated in whole blood of 93 adolescents in association with a modeled risk score for MDD. To the best of our knowledge, this is the first study investigating genome-wide DNA methylation shifts in relationship to MDD in a population-based cohort of

adolescents (Vania Januar, Saffery, & Ryan, 2015). We were able to validate our finding in an independent data set of 78 adolescents and could also show that that the methylation drift of the MIR4646 associated CpG locus was inversely correlated with the degree of gene expression in healthy individuals. Thus, we detect miR-4646 as a novel marker for risk of MDD in adolescence, a critical period for environmental experience, and could add value in planning the preventive and control strategies for adulthood MDD.

We detect hypomethylation of cg04102384 – a methylation locus located within 246 bp of the transcriptional start site for MIR4646 - and the risk of depression in adolescents. By further investigating associations between methylation and expression in healthy individuals, we provide evidence that the level of methylation at this CpG site regulates expression of MIR4646-3p in whole blood, underlying the functional importance of this locus. Our functional analysis revealed that the MIR4646 associated CpG site is located in an enhancer region of the gene throughout all investigated brain tissues, as well as in peripheral blood cells (see **Figure 3**). These homogenous findings for both brain tissue and blood are of importance as it underlines the applicability/transferability of epigenetic findings made in blood on changes putatively also abundant in brain in addition to the correlations between blood and brain methylomes(Hannon, Lunnon, Schalkwyk, & Mill, 2015; Nikolova et al., 2014). Moreover, in the MDD context, the whole blood methylation analysis may be relevant (Walker et al., 2016) particularly for the multiple pathophysiological pathways associated with this disease, e.g. inflammatory (Zunszain, Hepgul, & Pariante, 2013), immune (Blume, Douglas, & Evans, 2011) and metabolic processes (Vogelzangs et al., 2014).

Strikingly, MIR4646-3p appears to play a role in the neurobiological context, as detected by KEGG pathway annotation. In our study, we identify that MIR4646-3p is associated with the elongation of fatty acids. Importantly, this process mediates the conversion of omega-3 fatty acids, e.g. the eicosapentaenoic (EPA) and docosahexaenoic (DHA). Essential omega-3 fatty

acids cannot be synthesized by the organism and have to be taken up by the diet, from sources as fish, eggs and soy products. They were shown to be involved in brain functioning, particularly in memory and cognitive performance (Ruxton, Reed, Simpson, & Millington, 2004). Interestingly, omega-3 fatty acids appear to be implicated in MDD pathophysiology (Hibbeln, 1998; Logan, 2004; Mamalakis, Tornaritis, & Kafatos, 2002; Peet, Murphy, Shay, & Horrobin, 1998). A study showed that pretreated mice with fish oil have developed less depressive symptoms through the suppression of neuroinflammation (Numata et al., 2015). Grayson et al. revealed in neuroimaging studies that lack of omega-3 implicates functional dysregulation at prefrontal cortical connectivity in monkeys (Grayson, Kroenke, Neuringer, & Fair, 2014). Notably, while it was shown that adolescents with MDD have a decreased frontal white matter integrity and worse connectivity within frontal lobe cortical networks (Connolly et al., 2013; Ho et al., 2014), Chhetry et al. found that fish oil supplementation could improve the connectivity, resulting in a reduction in depression symptom severity (Chhetry et al., 2016). The global importance of fatty acid elongation leaves of course space for additional regulatory implications in MDD. An unbalanced ratio of omega-6/omega-3 could contribute to obesity (Simopoulos, 2016), which has been demonstrated to increase the risk of depression (Luppino et al., 2010). Previously, MIR4646 was identified to be upregulated in peripheral blood from schizophrenic patients (Fan et al., 2015). Interestingly, several studies described deficiencies of omega-3 fatty acids in red blood cells membranes from schizophrenic patients (Peet, Shah, Selvam, & Ramchand, 2004; Reddy, Keshavan, & Yao, 2004). These common findings regarding the omega-3 role through the MIR4646 suggest a partially shared etiology of MDD and schizophrenia.

No study previously described the importance of MIR4646. The conservation throughout miRNA-4646 stem-loop region and in particular the seed regions of the 5p and 3p mature sequences, along with pre-microRNA folding structure predictions in all of the investigated

sequences and the conserved gene synteny suggests that this microRNA region is conserved from advent of the Simiiformes (monkey). The alignment of the reverse complement of miRNA-4646-3p to the miRNA-204 family is interesting as the reverse complement of miRNAs are speculated to be able to work in subtle regulatory activity and have some functional activity (Shao, Ma, Xu, Wang, & Meng, 2012). Additionally, the mir-204 family has been investigated in association with schizophrenia and regulating non-coding RNA that affect neurotransmitter and ion channel gene sets (Cammaerts et al., 2015).

Previous genome-wide studies showed evidence regarding altered DNA methylation in depression disorder, but mainly in monozygotic twin pairs, which are not reflecting the normal population (Byrne et al., 2013; Cordova-Palomera et al., 2015; Dempster et al., 2014). In our study, we could take into account both genetics and environmental components as an interconnected biological network, by investigating a population-based cohort. Moreover, the advantages of studying an adolescent population at risk of MDD instead of investigating individuals showing already clinical symptoms are numerous. Firstly, as a childhood episode of MDD substantially increases the risk of adverse outcomes in adulthood (Fergusson, Boden, & Horwood, 2007), it calls attention for early detection of MDD as a marker for severity of the underlying vulnerability. Studying a cohort at risk for MDD may thus provide an excellent platform for identifying epigenetic susceptibility factors that hold true for depressive disorders in general. Secondly, it can be argued that investigating developmental stages of psychopathology allows better isolating of causative factors for disease pathology, whereas analyses performed on adults that have had psychiatric problems for a long time are more likely to be burdened by confounding factors, e.g. antidepressants, alcohol or substance abuse, smoking, cardiovascular complications or even aging, all of which may influence methylation and could produce false outcomes (Breitling, Yang, Korn, Burwinkel, & Brenner, 2011; Menke & Binder, 2014; Zampieri et al., 2015).

While DNA methylation is the best characterized mechanism of epigenetic regulation (Novik et al., 2002), different epigenetic mechanisms may act co-dependently and participate in cross-regulating activities (Feil & Fraga, 2012; Iorio, Piovan, & Croce, 2010; Sun, Kennedy, & Nestler, 2013). Klengel et a. suggested that a better understanding of DNA methylation and its link to other epigenetic mechanisms will be important for understanding the epigenetic contribution to psychopathologies (Klengel, Pape, Binder, & Mehta, 2014). In this respect, our findings provide further grist to the mill that the epigenetic contribution to psychopathologies involve co-dependent and cross-regulatory processes of DNA methylation and miRNA activity. The DAWBA bands have been shown to be similar or identical to clinician-rated diagnoses in estimated effect sizes, significance levels and substantive conclusions regarding risk factor associations (Goodman et al., 2011). As the scores were not obtained in a face-to-face meeting of a clinician and the child, however, it cannot be completely excluded that an individual subject received an unrepresentative risk score assessment. Our study is also restricted by the limited number of individuals with higher risk-score assessments of MDD in both discovery (n=22) and replication group (n=13). Future studies including a high number of individuals would be of great value to confirm the results presented. Furthermore, in the replication data set, there were between-group differences in BMI z-score defined weight categories and in the use of contraceptive pills. Due to power issues, we were not able to adjust for these factors on the confirmatory association analysis between DNA methylation at cg04102384 and depression risk group. As such, this represents a limitation which could be a source of potential bias. In addition, the putative gene targets of MIR4646 were identified by computer algorithms. In vivo studies are needed to confirm the true gene targets of MIR4646 and its relation to the fatty acid elongation pathway. In addition, we were not able to confirm MIR4646 to be differentially expressed between adolescents at high risk of depression and controls in the qRT-PCR measurement. A possible reason could be the low fold change found of this miRNA levels

between cases and controls in the discovery cohort (~21,4%) and the small sample size (n=27), an observation which is in line with another study that could not validate miRNA levels with less than four-fold change of expression (Fan et al., 2015). This limitation of qRT-PCR measurements could also contribute to explaining our inability to confirm the observed inverse correlation between methylation and expression at MIR4646-3p. Future studies should investigate to what extent miR-4646 is differentially expressed in individuals suffering from depression and its relation to the identified methylation locus.

In summary we were able to reveal and validate a hypomethylated CpG site within MIR4646 significantly associated with a modulated risk score for MDD in 223 adolescents. Notably, our study shows evidence that the DNA methylation changes of cg04102384 could regulate the gene expression pattern of MIR4646-3p. The performed functional analysis highlights a potential role of this miRNA in the regulation of fatty acids, with potential involvement in the conversion of omega-3 fatty acids. Our findings indicate that the dynamic interplay between DNA methylation and miRNA-transcription could have an elucidative mechanism on the pathogenesis of MDD and related diseases.

Acknowledgement: Methylation profiling was performed by the SNP&SEQ Technology Platform in Uppsala. The platform is part of Science for Life Laboratory at Uppsala University and supported as a national infrastructure by the Swedish Research Council. DNA extraction was performed by Latvian BMC, funded by the Latvian Council of Science European Social Fund and European Regional Development Fund. The studies were supported by the Swedish Research Foundation and the Swedish Brain Foundation.

References

Aryee, M. J., Jaffe, A. E., Corrada-Bravo, H., Ladd-Acosta, C., Feinberg, A. P., Hansen, K. D., & Irizarry, R. A. (2014). Minfi: A flexible and comprehensive Bioconductor package for the analysis of Infinium DNA methylation microarrays. *Bioinformatics*, *30*(10), 1363–1369. https://doi.org/10.1093/bioinformatics/btu049

Aulchenko, Y. S., Ripke, S., Isaacs, A., & van Duijn, C. M. (2007). GenABEL: An R library for genome-wide association analysis. *Bioinformatics*, *23*(10), 1294–1296. https://doi.org/10.1093/bioinformatics/btm108

Betel, D., Koppal, A., Agius, P., Sander, C., & Leslie, C. (2010). Comprehensive modeling of microRNA targets predicts functional non-conserved and non-canonical sites. *Genome Biology*, *11*(8), R90. https://doi.org/10.1186/gb-2010-11-8-r90

Birmaher, B., Arbelaez, C., & Brent, D. (2002). Course and outcome of child and adolescent major depressive disorder. *Child Adolesc Psychiatr Clin N Am*, *11*(3), 619–37, x. https://doi.org/10.1016/S1056-4993(02)00011-1

Blume, J., Douglas, S. D., & Evans, D. L. (2011). Immune suppression and immune activation in depression. *Brain, Behavior, and Immunity*. https://doi.org/10.1016/j.bbi.2010.10.008

Boratyn, G. M., Camacho, C., Cooper, P. S., Coulouris, G., Fong, A., Ma, N., … Zaretskaya, I. (2013). BLAST: a more efficient report with usability improvements. *Nucleic Acids Research*, *41*(Web Server issue). https://doi.org/10.1093/nar/gkt282

Breitling, L. P., Yang, R., Korn, B., Burwinkel, B., & Brenner, H. (2011). Tobacco-smoking-related differential DNA methylation: 27K discovery and replication. *American Journal of Human Genetics*, *88*(4), 450–457. https://doi.org/10.1016/j.ajhg.2011.03.003

Byrne, E. M., Carrillo-Roa, T., Henders, A. K., Bowdler, L., McRae, A. F., Heath, A. C., … Wray, N. R. (2013). Monozygotic twins affected with major depressive disorder have greater variance in methylation than their unaffected co-twin. *Transl Psychiatry*, *3*, e269. https://doi.org/10.1038/tp.2013.45

Cammaerts, S., Strazisar, M., Smets, B., Weckhuysen, S., Nordin, A., De Jonghe, P., … Del Favero, J. (2015). Schizophrenia-Associated MIR204 Regulates Noncoding RNAs and Affects Neurotransmitter and Ion Channel Gene Sets. *PLoS ONE*, *10*(12). https://doi.org/10.1371/journal.pone.0144428

Chen, W., Gao, G., Nerella, S., Hultman, C. M., Magnusson, P. K. E., Sullivan, P. F., … van den Oord, E. J. C. G. (2013). MethylPCA: a toolkit to control for confounders in methylome-wide association studies. *BMC Bioinformatics*, *14*, 74. https://doi.org/10.1186/1471-2105-14-74

Chen, W., & Qin, C. (2015). General hallmarks of microRNAs in brain evolution and development. *RNA Biology*, *12*(7), 701–708. https://doi.org/10.1080/15476286.2015.1048954

Chhetry, B. T., Hezghia, A., Miller, J. M., Lee, S., Rubin-Falcone, H., Cooper, T. B., … Sublette, M.

E. (2016). Omega-3 polyunsaturated fatty acid supplementation and white matter changes in major depression. *Journal of Psychiatric Research*, *75*, 65–74. https://doi.org/10.1016/j.jpsychires.2015.12.007

Connolly, C. G., Wu, J., Ho, T. C., Hoeft, F., Wolkowitz, O., Eisendrath, S., … Yang, T. T. (2013). Resting-state functional connectivity of subgenual anterior cingulate cortex in depressed adolescents. *Biological Psychiatry*, *74*(12), 898–907. https://doi.org/10.1016/j.biopsych.2013.05.036

Cordova-Palomera, A., Fatjo-Vilas, M., Gasto, C., Navarro, V., Krebs, M. O., & Fananas, L. (2015). Genome-wide methylation study on depression: differential methylation and variable methylation in monozygotic twins. *Transl Psychiatry*, *5*, e557. https://doi.org/10.1038/tp.2015.49

Costello, E. J., Erkanli, A., & Angold, A. (2006). Is there an epidemic of child or adolescent depression? *Journal of Child Psychology and Psychiatry*, *47*(12), 1263–1271. https://doi.org/Doi 10.1111/J.1469-7610.2006.01682.X

D'Addario, C., Dell'Osso, B., Palazzo, M. C., Benatti, B., Lietti, L., Cattaneo, E., … Altamura, A. C. (2012). Selective DNA Methylation of BDNF Promoter in Bipolar Disorder: Differences Among Patients with BDI and BDII. *Neuropsychopharmacology*, *37*(7), 1647–1655. https://doi.org/10.1038/npp.2012.10

Dell'Osso, B., D'Addario, C., Carlotta Palazzo, M., Benatti, B., Camuri, G., Galimberti, D., … Carlo Altamura, A. (2014). Epigenetic modulation of BDNF gene: Differences in DNA methylation between unipolar and bipolar patients. *Journal of Affective Disorders*, *166*, 330–333. https://doi.org/10.1016/j.jad.2014.05.020

Dempster, E. L., Wong, C. C. Y., Lester, K. J., Burrage, J., Gregory, A. M., Mill, J., & Eley, T. C. (2014). Genome-wide methylomic analysis of monozygotic twins discordant for adolescent depression. *Biological Psychiatry*, *76*(12), 977–983. https://doi.org/10.1016/j.biopsych.2014.04.013

Du, P., Zhang, X., Huang, C.-C., Jafari, N., Kibbe, W. A., Hou, L., & Lin, S. M. (2010). Comparison of Beta-value and M-value methods for quantifying methylation levels by microarray analysis. *BMC Bioinformatics*, *11*(1), 587. https://doi.org/10.1186/1471-2105-11-587

Dweep, H., Sticht, C., Pandey, P., & Gretz, N. (2011). MiRWalk - Database: Prediction of possible miRNA binding sites by " walking" the genes of three genomes. *Journal of Biomedical Informatics*, *44*(5), 839–847. https://doi.org/10.1016/j.jbi.2011.05.002

Fan, H.-M., Sun, X.-Y., Niu, W., Zhao, L., Zhang, Q.-L., Li, W.-S., … Lu, J. (2015). Altered microRNA Expression in Peripheral Blood Mononuclear Cells from Young Patients with Schizophrenia. *Journal of Molecular Neuroscience : MN*, *56*(3), 562–571. https://doi.org/10.1007/s12031-015-0503-z

Feil, R., & Fraga, M. F. (2012). Epigenetics and the environment: emerging patterns and implications. *Nature Reviews Genetics*, *13*(2), 97–109. https://doi.org/10.1038/nrg3142

Fergusson, D. M., Boden, J. M., & Horwood, L. J. (2007). Recurrence of major depression in adolescence and early adulthood, and later mental health, educational and economic outcomes. *The British Journal of Psychiatry : The Journal of Mental Science*, *191*(4), 335–42. https://doi.org/10.1192/bjp.bp.107.036079

Glied, S., & Pine, D. S. (2002). Consequences and correlates of adolescent depression. *Archives of Pediatrics & Adolescent Medicine*, *156*(10), 1009–14. https://doi.org/10.1001/archpedi.156.10.1009

Goodman, A., Heiervang, E., Collishaw, S., & Goodman, R. (2011). The "DAWBA bands" as an ordered-categorical measure of child mental health: Description and validation in British and Norwegian samples. *Social Psychiatry and Psychiatric Epidemiology*, *46*(6), 521–532. https://doi.org/10.1007/s00127-010-0219-x

Grayson, D. S., Kroenke, C. D., Neuringer, M., & Fair, D. A. (2014). Dietary omega-3 fatty acids modulate large-scale systems organization in the rhesus macaque brain. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *34*(6), 2065–74. https://doi.org/10.1523/JNEUROSCI.3038-13.2014

Grimson, A., Farh, K. K. H., Johnston, W. K., Garrett-Engele, P., Lim, L. P., & Bartel, D. P. (2007). MicroRNA Targeting Specificity in Mammals: Determinants beyond Seed Pairing. *Molecular Cell*, *27*(1), 91–105. https://doi.org/10.1016/j.molcel.2007.06.017

Hannon, E., Lunnon, K., Schalkwyk, L., & Mill, J. (2015). Interindividual methylomic variation

across blood, cortex, and cerebellum: Implications for epigenetic studies of neurological and neuropsychiatric phenotypes. *Epigenetics*, *10*(11), 1024–1032. https://doi.org/10.1080/15592294.2015.1100786

Hibbeln, J. R. (1998). Fish consumption and major depression. *Lancet*, *351*(9110), 1213. https://doi.org/10.1016/S0140-6736(05)79168-6

Hinrichs, A. L., Larkin, E. K., & Suarez, B. K. (2009). Population stratification and patterns of linkage disequilibrium. In *Genetic Epidemiology* (Vol. 33). https://doi.org/10.1002/gepi.20478

Ho, T. C., Yang, G., Wu, J., Cassey, P., Brown, S. D., Hoang, N., … Yang, T. T. (2014). Functional connectivity of negative emotional processing in adolescent depression. *Journal of Affective Disorders*, *155*(1), 65–74. https://doi.org/10.1016/j.jad.2013.10.025

Iorio, M. V., Piovan, C., & Croce, C. M. (2010). Interplay between microRNAs and the epigenetic machinery: An intricate network. *Biochimica et Biophysica Acta - Gene Regulatory Mechanisms*. https://doi.org/10.1016/j.bbagrm.2010.05.005

Jansen, R., Penninx, B. W. J. H., Madar, V., Xia, K., Milaneschi, Y., Hottenga, J. J., … Sullivan, P. F. (2015). Gene expression in major depressive disorder. *Molecular Psychiatry*, (October 2014), 1–9. https://doi.org/10.1038/mp.2015.57

Januar, V., Ancelin, M. L., Ritchie, K., Saffery, R., & Ryan, J. (2015). BDNF promoter methylation and genetic variation in late-life depression. *Transl Psychiatry*, *5*, e619. https://doi.org/10.1038/tp.2015.114

Januar, V., Saffery, R., & Ryan, J. (2015). Epigenetics and depressive disorders: A review of current progress and future directions. *International Journal of Epidemiology*, *44*(4), 1364–1387. https://doi.org/10.1093/ije/dyu273

Kamburov, A., Stelzl, U., Lehrach, H., & Herwig, R. (2013). The ConsensusPathDB interaction database: 2013 Update. *Nucleic Acids Research*, *41*(D1). https://doi.org/10.1093/nar/gks1055

Katoh, K., & Standley, D. M. (2013). MAFFT multiple sequence alignment software version 7: Improvements in performance and usability. *Molecular Biology and Evolution*, *30*(4), 772–780. https://doi.org/10.1093/molbev/mst010

Klengel, T., Pape, J., Binder, E. B., & Mehta, D. (2014). The role of DNA methylation in stress-

related psychiatric disorders. *Neuropharmacology*. https://doi.org/10.1016/j.neuropharm.2014.01.013

Kozomara, A., & Griffiths-Jones, S. (2014). MiRBase: Annotating high confidence microRNAs using deep sequencing data. *Nucleic Acids Research*, *42*(D1). https://doi.org/10.1093/nar/gkt1181

Lê, S., Josse, J., & Mazet, F. (2008). Package " FactoMineR ." *Journal of Statistical Software*, *25*(1), 1–18. https://doi.org/10.1007/978-3-540-74686-7

Leek, J. T., Johnson, W. E., Parker, H. S., Jaffe, A. E., & Storey, J. D. (2012). The SVA package for removing batch effects and other unwanted variation in high-throughput experiments. *Bioinformatics*, *28*(6), 882–883. https://doi.org/10.1093/bioinformatics/bts034

Letunic, I., & Bork, P. (2016). Interactive tree of life (iTOL) v3: an online tool for the display and annotation of phylogenetic and other trees. *Nucleic Acids Research*. https://doi.org/10.1093/nar/gkw290

Logan, A. C. (2004). Omega-3 fatty acids and major depression: a primer for the mental health professional. *Lipids in Health and Disease*, *3*(1), 25. https://doi.org/10.1186/1476-511X-3-25

Luppino, F. S., de Wit, L. M., Bouvy, P. F., Stijnen, T., Cuijpers, P., Penninx, B. W. J. H., & Zitman, F. G. (2010). Overweight, obesity, and depression: a systematic review and meta-analysis of longitudinal studies. *Archives of General Psychiatry*, *67*(3), 220–9. https://doi.org/10.1001/archgenpsychiatry.2010.2

Mamalakis, G., Tornaritis, M., & Kafatos, a. (2002). Depression and adipose essential polyunsaturated fatty acids. *Prostaglandins, Leukotrienes, and Essential Fatty Acids*, *67*(5), 311–318. https://doi.org/10.1054/plef.2002.0435

Maunakea, A. K., Nagarajan, R. P., Bilenky, M., Ballinger, T. J., D'Souza, C., Fouse, S. D., … Costello, J. F. (2010). Conserved role of intragenic DNA methylation in regulating alternative promoters. *Nature*, *466*(7303), 253–7. https://doi.org/10.1038/nature09165

Menke, A., & Binder, E. B. (2014). Epigenetic alterations in depression and antidepressant treatment. *Dialogues in Clinical Neuroscience*, *16*(3), 395–404. https://doi.org/10.1097/gme.0b013e3181e15e0e

Morris, T. J., Butcher, L. M., Feber, A., Teschendorff, A. E., Chakravarthy, A. R., Wojdacz, T. K., &

Beck, S. (2014). ChAMP: 450k Chip Analysis Methylation Pipeline. *Bioinformatics*, *30*(3), 428–430. https://doi.org/10.1093/bioinformatics/btt684

Nikolova, Y. S., Koenen, K. C., Galea, S., Wang, C.-M., Seney, M. L., Sibille, E., … Hariri, A. R. (2014). Beyond genotype: serotonin transporter epigenetic modification predicts human brain function. *Nature Neuroscience*, *17*(9), 1153–5. https://doi.org/10.1038/nn.3778

Novik, K. L., Nimmrich, I., Genc, B., Maier, S., Piepenbrock, C., Olek, A., & Beck, S. (2002). Epigenomics: Genome-wide study of methylation phenomena. *Current Issues in Molecular Biology*, *4*(4), 111–128.

Nugent, N. R., Tyrka, A. R., Carpenter, L. L., & Price, L. H. (2011). Gene-environment interactions: Early life stress and risk for depressive and anxiety disorders. *Psychopharmacology*. https://doi.org/10.1007/s00213-010-2151-x

Numata, S., Ishii, K., Tajima, A., Iga, J., Kinoshita, M., Watanabe, S., … Ohmori, T. (2015). Blood diagnostic biomarkers for major depressive disorder using multiplex DNA methylation profiles: discovery and validation. *Epigenetics : Official Journal of the DNA Methylation Society*, *10*(2), 135–41. https://doi.org/10.1080/15592294.2014.1003743

O'Leary, N. A., Wright, M. W., Brister, J. R., Ciufo, S., Haddad, D., McVeigh, R., … Pruitt, K. D. (2016). Reference sequence (RefSeq) database at NCBI: Current status, taxonomic expansion, and functional annotation. *Nucleic Acids Research*, *44*(D1), D733–D745. https://doi.org/10.1093/nar/gkv1189

Olsson, C. A., Foley, D. L., Parkinson-Bates, M., Byrnes, G., McKenzie, M., Patton, G. C., … Saffery, R. (2010). Prospects for epigenetic research within cohort studies of psychological disorder: A pilot investigation of a peripheral cell marker of epigenetic risk for depression. *Biological Psychology*, *83*(2), 159–165. https://doi.org/10.1016/j.biopsycho.2009.12.003

Paraskevopoulou, M. D., Georgakilas, G., Kostoulas, N., Vlachos, I. S., Vergoulis, T., Reczko, M., … Hatzigeorgiou, A. G. (2013). DIANA-microT web server v5.0: service integration into miRNA functional analysis workflows. *Nucleic Acids Research*, *41*(Web Server issue). https://doi.org/10.1093/nar/gkt393

Peet, M., Murphy, B., Shay, J., & Horrobin, D. (1998). Depletion of omega-3 fatty acid levels in red

blood cell membranes of depressive patients. *Biological Psychiatry*, *43*(5), 315–9. https://doi.org/10.1016/S0006-3223(97)00206-0

Peet, M., Shah, S., Selvam, K., & Ramchand, C. N. (2004). Polyunsaturated fatty acid levels in red cell membranes of unmedicated schizophrenic patients. *World J Biol Psychiatry*, *5*(2), 92–99. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/15179668

Rask-Andersen, M., Bringeland, N., Nilsson, E. K., Bandstein, M., Búcaro, M. O., Vogel, H., … Schiöth, H. B. (2016). Postprandial alterations in whole-blood DNA methylation are mediated by changes in white blood cell composition. *American Journal of Clinical Nutrition*, *104*(2), 518–525. https://doi.org/10.3945/ajcn.115.122366

Reddy, R. D., Keshavan, M. S., & Yao, J. K. (2004). Reduced red blood cell membrane essential polyunsaturated fatty acids in first episode schizophrenia at neuroleptic-naive baseline. *Schizophrenia Bulletin*, *30*(4), 901–911.

Rehmsmeier, M., Steffen, P., Hochsmann, M., & Giegerich, R. (2004). Fast and effective prediction of microRNA/target duplexes. *RNA (New York, N.Y.)*, *10*(10), 1507–17. https://doi.org/10.1261/rna.5248604

Reinius, L. E., Acevedo, N., Joerink, M., Pershagen, G., Dahlen, S. E., Greco, D., … Kere, J. (2012). Differential DNA methylation in purified human blood cells: implications for cell lineage and studies on disease susceptibility. *PLoS One*, *7*(7), e41361. https://doi.org/10.1371/journal.pone.0041361

Ruxton, C. H. S., Reed, S. C., Simpson, M. J. A., & Millington, K. J. (2004). The health benefits of omega-3 polyunsaturated fatty acids: A review of the evidence. *Journal of Human Nutrition and Dietetics*. https://doi.org/10.1111/j.1365-277X.2004.00552.x

Ruzicka, W. B., Subburaju, S., & Benes, F. M. (2015). Circuit- and Diagnosis-Specific DNA Methylation Changes at γ-Aminobutyric Acid–Related Genes in Postmortem Human Hippocampus in Schizophrenia and Bipolar Disorder. *JAMA Psychiatry*, *2478*(6), 541–551. https://doi.org/10.1001/jamapsychiatry.2015.49

Sabunciyan, S., Aryee, M. J., Irizarry, R. A., Rongione, M., Webster, M. J., Kaufman, W. E., … Gen, R. E. D. C. (2012). Genome-wide DNA methylation scan in major depressive disorder. *PLOS*

One, *7*(4), e34451. https://doi.org/10.1371/journal.pone.0034451

Sambrook, J., Fritsch, E. F., & Maniatis, T. (1989). *MOLECULAR CLONING A LABORATORY MANUAL SECOND EDITION VOLS. 1 2 AND 3*. *Sambrook J E F Fritsch and T Maniatis Molecular Cloning A Laboratory Manual Second Edition Vols 1 2 and 3 XxxixPagination VariesVol 1 XxxiiiPagination VariesVol 2 XxxiiPagination VariesVol 3 Cold Spring Harbor Laboratory Press*.

Schalkwyk, L. C., Pidsley, R., & Wong, C. C. Y. (2013). wateRmelon: Illumina 450 methylation array normalization and metrics. *R Package Version 1.2.2*.

Schmitt, A., Malchow, B., Hasan, A., & Falkai, P. (2014). The impact of environmental factors in severe psychiatric disorders. *Frontiers in Neuroscience*, *8*, 19. https://doi.org/10.3389/fnins.2014.00019; 10.3389/fnins.2014.00019

Shao, C., Ma, X., Xu, X., Wang, H., & Meng, Y. (2012). Genome-Wide Identification of Reverse Complementary microRNA Genes in Plants. *PLoS ONE*, *7*(10). https://doi.org/10.1371/journal.pone.0046991

Simopoulos, A. P. (2016). An increase in the Omega-6/Omega-3 fatty acid ratio increases the risk for obesity. *Nutrients*. https://doi.org/10.3390/nu8030128

Smyth, G. K. (2004). Linear models and empirical bayes methods for assessing differential expression in microarray experiments. *Statistical Applications in Genetics and Molecular Biology*, *3*, Article3. https://doi.org/10.2202/1544-6115.1027

Sun, H., Kennedy, P. J., & Nestler, E. J. (2013). Epigenetics of the depressed brain: role of histone acetylation and methylation. *Neuropsychopharmacology : Official Publication of the American College of Neuropsychopharmacology*, *38*(1), 124–37. https://doi.org/10.1038/npp.2012.73

Villela, D., Ramalho, R. F., Silva, A. R. T., Brentani, H., Suemoto, C. K., Pasqualucci, C. A., … Rosenberg, C. (2016). Differential DNA Methylation of MicroRNA Genes in Temporal Cortex from Alzheimer's Disease Individuals. *Neural Plasticity*, *2016*. https://doi.org/10.1155/2016/2584940

Vogelzangs, N., Beekman, A. T. F., van Reedt Dortland, A. K. B., Schoevers, R. A., Giltay, E. J., de Jonge, P., & Penninx, B. W. J. H. (2014). Inflammatory and metabolic dysregulation and the 2-

year course of depressive disorders in antidepressant users. *Neuropsychopharmacology : Official Publication of the American College of Neuropsychopharmacology*, *39*(7), 1624–34. https://doi.org/10.1038/npp.2014.9

Walker, R. M., Sussmann, J. E., Whalley, H. C., Ryan, N. M., Porteous, D. J., McIntosh, A. M., & Evans, K. L. (2016). Preliminary assessment of pre-morbid DNA methylation in individuals at high genetic risk of mood disorders. *Bipolar Disorders*, *18*(5), 410–422. https://doi.org/10.1111/bdi.12415

Waterhouse, A. M., Procter, J. B., Martin, D. M. A., Clamp, M., & Barton, G. J. (2009). Jalview version 2: A Multiple Sequence Alignment and Analysis Workbench. *Bioinformatics*, *25*(9), 1189–1191. https://doi.org/10.1093/bioinformatics/btp033 [doi]

Zaghlool, S. B., Al-Shafai, M., Al Muftah, W. A., Kumar, P., Falchi, M., & Suhre, K. (2015). Association of DNA methylation with age, gender, and smoking in an Arab population. *Clinical Epigenetics*, *7*(6), 1–12. https://doi.org/10.1186/s13148-014-0040-6

Zampieri, M., Ciccarone, F., Calabrese, R., Franceschi, C., B??rkle, A., & Caiafa, P. (2015). Reconfiguration of DNA methylation in aging. *Mechanisms of Ageing and Development*. https://doi.org/10.1016/j.mad.2015.02.002

Zunszain, P. A., Hepgul, N., & Pariante, C. M. (2013). Inflammation and Depression. *Current Topics in Behavioral Neurosciences*, *14*, 135–151. https://doi.org/10.1007/7854_2012_211